Title
Before and After Photoshop: Recursive Fraud in the Age of Digital Reproducibility.

Journal
Angewandte Chemie (International ed. in English), 58(46)

ISSN
1433-7851

Author
Biagioli, Mario

Publication Date
2019-11-01

DOI
10.1002/anie.201908646

Peer reviewed

International Edition: DOI: 10.1002/anie.201908646
German Edition: DOI: 10.1002/ange.201908646

Before and After Photoshop: Recursive Fraud in the Age of Digital Reproducibility**

*Mario Biagioli**

Mario Biagioli

Keywords: fraud · images · manipulation · metrics · misconduct

Concerns with scientific misconduct are typically expressed in quantitative terms: How frequent is misconduct? Does an increase in retractions signal an increase in misconduct or just improved detection? What is the geographical distribution of fraud? These are important questions, but it is equally important to realize that misconduct is undergoing striking qualitative changes as well. Some scientists used to fabricate and falsify claims and evidence, but now they also aim at manipulating the publication process to enhance their performance metrics. Examples include fake peer reviews that facilitate publication in higher impact journals, citation clubs and rings, or the hacking of the peer review process to request more citations to one's work in return for a positive review of the manuscript. Because they typically manipulate the publication and reception of the work rather than its research and writing, I call these practices "postproduction misconduct."[1]

This does not mean that traditional misconduct—fabrication, falsification, and plagiarism—has been rendered obsolete by these new trends. To the contrary, the pressure to meet performance targets combined with the availability of new digital editing and publishing tools has brought significant innovation to traditional misconduct. With the widespread online availability of texts and word-processing software that made cut-and-paste practices easier and faster, plagiarism was the first type of misconduct to enter the digital mode of production. Then in 2005 three MIT graduate students released SCIgen, an open-source software generating fake but plausibly sounding computer science articles.[2] While explicitly conceived as a prank or hoax to expose the low selectivity of some journals and conferences, SCIgen was instead quickly adopted by fraudsters eager to use it for real. Years later, Springer and IEEE had to retract more than 120 gibberish papers and conference abstracts that had slipped through their journals' review.[3]

The most conspicuous example of digital misconduct, however, comes from Photoshop and other software that can turn image manipulation into a labor-saving, computer-aided form of scientific fraud. Most cases concern western blots, but any kind of photographic evidence (including other types of blots, micrographs, TEM images) may be involved. Photoshop makes it possible to crop, delete, splice, compress, stretch, tilt, flip, and recombine bands, modify their contrast level to produce fraudulent but professional-looking images. These images may then be used as evidence not just in one single article but in several publications, recycling the same manipulated blot to falsely reference different experimental conditions and samples. Two recent studies report that between 3.8% and 6% of the articles sampled contained duplicated and/or manipulated images.[4] In addition, about 60% of the 2018 misconduct findings by the US Office or Research Integrity involved image manipulation. These are remarkable trends given that such manipulations have started only in the mid 1990s.

Word-processing software, SCIgen, or Photoshop have increased fraudulent output, but—and this is my claim here—they have also changed the very form of fraud. Older fraud involved unique, local, hand-manipulated evidence crafted to support specific claims. The 1912 Piltdown Man case involved planting physical evidence (medieval human craniums mixed with modern young orangutan jawbones) that were physically manipulated and placed in a pit to create evidence of the "missing link" between humans and apes. Similarly, in 1926 Paul Kammerer was accused of physically tattooing "nuptial pads" on midwife toads in his laboratory in Vienna to support Lamarckian against Darwinian evolution, and in 1974 William Summerlin admitted to using a felt-tip pen to color a skin graft on an albino mouse at Sloan-Kettering in New York to claim that he had been able to suppress immunological rejection of transplanted tissues. All these cases are essentially different from today's digitally manipulated images, SCIgen-produced articles, or electronically plagiarized texts. The "crime scene" has been displaced

[*] Prof. Dr. M. Biagioli
University of California Los Angeles, Dept of Communication and School of Law, UCLA School of Law
385 Charles E. Young Drive East, 1242 Law Building
Los Angeles, CA 90095 (USA)
E-mail: mbiagioli@ucdavis.edu

[**] Research for this Editorial was partly supported by the DFG Centre of Advanced Studies "Imaginaria of Force", University of Hamburg.

The ORCID identification number for the author of this article can be found under: https://doi.org/10.1002/anie.201908646.

from the field or laboratory to the computer, and while the mode of production remains artisanal, digital tools allow individual fraudsters to achieve quasi-industrial productivity.

It is important to realize, however, that digital tools make the *reproduction* of fraud—not just its production—virtually effortless. An example is the recycling of identical or slightly modified fraudulent blots to support different claims in different articles. In one instance, the same blot was identified in fifteen images across ten different articles.[5] Photoshop and similar tools are thus enabling something that is categorically new: *recursive fraud*, that is, the serial digital reproduction and dissemination of more fraud based on tweaks or copies of previous fraudulent evidence. Perhaps perversely, one could compare this transformation to the impact of the introduction of photography on the status of previously unique artworks.

But if some of the new fraud has become derivative, does that mean that traditional fraud was authentic? In a strange sense it was. An art forger who paints a "lost Rembrandt" will try to pass it off as one authored by the Dutch master, which it is not. Still, that forgery remains a unique painting original to the forger. Similarly, Summerlin's painted mice may not have been works of art, but they were still hand-made originals, not copies. They were *fraudulent and yet authentic.* By contrast, today's serially photoshopped images are *fraudulent copies.* They are re-manipulations of previous manipulations or, in a few cases, non-fraudulent images that are copied and re-purposed for fraudulent ends.

Technologies that make it possible to move from producing singular authentic fraud to disseminating multiple fraudulent copies is precisely what makes the new digital fraud so productive. And that brings us back to metrics. If it were not for an obsession with productivity, why would we find the same fraudulent blot reproduced, more or less identically, fifteen times in ten publications? Or why would we see a photo of damage produced by a tapeworm infection re-published verbatim (after being plagiarized from another article) as an image of spinal cord compression? Or a photograph of human brain tissue that turns out to be an image of mouse brain tissue lifted from a public website?[6] The fraudsters seems to be in such a rush that they take shortcuts *even within fraud-making*, re-purposing evidence or reprinting it verbatim with a different caption because that is much faster than faking it from scratch, as fraudsters did in the pre-digital age.

An extensive study of image duplication and manipulation in biomedical publications has detected an inverse correlation between the frequency of problematic papers and the impact factor of the journals in which they are published. To the authors this suggests "that higher impact journals are better able to detect anomalous images prior to publication."[4a] It is certainly possible that, as the authors suggest, journals with more resources and more reputation to lose are better at catching duplicate and/or fraudulent images. But it is equally possible that high-impact journals receive fewer submissions containing manipulated images because their authors may choose lower impact venues to reduce the chance of detection. (SciGen articles and plagiarized publications also tend to gravitate to second- or third-tier venues.)

The best advantage that digitally manipulated images offer to fraudsters is that they look professional and thus prima facie credible. But that thin veneer of credibility lasts only until readers or editors probe a little harder. People with good eyes and some training can spot duplication and manipulations surprisingly well, and there is now software to detect more sophisticated manipulations. Similarly, passing off an image of a mouse brain for that or a human brain may work only until somebody familiar with any of those two images raises a red flag. Unlike traditional fraud where forensics could be lengthy and complex, digital fraud is likely to crumble like a house of cards shortly after one blows the whistle, as it often happens in places like PubPeer.

What we tend to see today are not ambitious fraudulent publications that try to escape detection thanks to their sophistication—the "perfect crime" scenario. The recursive fraud discussed here is closer to making many sloppy knock-offs of Levi's blue jeans than one masterful forgery of Leonardo. In the age of metrics, quantity trumps quality even in fraud. Rather than investing time and effort in developing hard-to-detect fabrications of important claims, fraudsters use standard digital tools to produce many articles that, while not likely to withstand much critical scrutiny, may survive in lower-tier journals where editors are less exacting and reading is optional. How can we otherwise explain the apparent carelessness of much of the fraudulent manipulation and reuse of blots or, even more strikingly, the reuse of plagiarized images that have nothing to do with the claim they are supposed to document? Rather than treating fraud as one thing—the production of falsehood—I suggest that we might want to think about its different economies, and the different forms that go with them.

[1] "Watch Out for Cheats in Citation Game": M. Biagioli, *Nature* **2016**, *535*, 201.
[2] https://pdos.csail.mit.edu/archive/scigen/.
[3] "Publishers withdraw more than 120 gibberish papers: Conference proceedings removed from subscription databases after scientist reveals that they were computer-generated" R. Van Noorden, *Nature*, 24 February **2014**.
[4] a) E. M. Bik, A. Casadevall, F. C. Fang, *mBio* **2016**, *7*, e00809–e00816; b) E. Bucci, *Cell Death Disease* **2018**, *9*, 400.
[5] L. Schneider, "Mario Saad and the Return of the Wandering Western Blot", *For Better Science*, June 21, **2016**, at https://forbetterscience.com/2016/06/21/mario-saad-and-the-return-of-the-wandering-western-blot/ (last accessed July 7, **2019**).
[6] V. Stern, "An author was accused of faking peer reviews. Turns out he also falsified two images", Retraction Watch, April 11, **2018** at https://retractionwatch.com/2018/04/11/an-author-was-accused-of-faking-peer-reviews-turns-out-he-also-falsified-two-images/.

Dateiname: O908646e
Pfad: L:/daten/AutoPrint.%0/VCH/preview/temp/
Status Neusatz
Datum: 38 KW., 18. September 2019 (Mittwoch)
Pagina: 3
Seite: 2 te von 3
Umfang (Seiten): 3
Zeit: 8:41:11 Uhr

Angewandte
International Edition Chemie

Guest Editorial

Scientific Fraud

M. Biagioli* ■■■■–■■■■

Before and After Photoshop: Recursive Fraud in the Age of Digital Reproducibility

"Word-processing software, SCIgen, or Photoshop have increased fraudulent output, but … they have also changed the very form of fraud. Photoshop and similar tools are thus enabling something that is categorically new: recursive fraud, that is, the serial digital reproduction and dissemination of more fraud."
Read more in the Guest Editorial by M. Biagioli.

SPACE RESERVED FOR IMAGE AND LINK

Share your work on social media! *Angewandte Chemie* has added Twitter as a means to promote your article. Twitter is an online microblogging service that enables its users to send and read short messages and media, known as tweets. Please check the pre-written tweet in the galley proofs for accuracy. If you, your team, or institution have a Twitter account, please include its handle @username. Please use hashtags only for the most important keywords, such as #catalysis, #nanoparticles, or #proteindesign. The ToC picture and a link to your article will be added automatically, so the **tweet text must not exceed 250 characters**. This tweet will be posted on the journal's Twitter account (follow us @angew_chem) upon publication of your article in its final (possibly unpaginated) form. We recommend you to re-tweet it to alert more researchers about your publication, or to point it out to your institution's social media team.

Please check that the ORCID identifiers listed below are correct. We encourage all authors to provide an ORCID identifier for each coauthor. ORCID is a registry that provides researchers with a unique digital identifier. Some funding agencies recommend or even require the inclusion of ORCID IDs in all published articles, and authors should consult their funding agency guidelines for details. Registration is easy and free; for further information, see http://orcid.org/.

Prof. Dr. Mario Biagioli http://orcid.org/0000-0003-0271-2979

Dateiname: O908646e
Pfad: L:/daten/AutoPrint.%0/VCH/preview/temp/
Status Neusatz
Datum: 38 KW., 18. September 2019 (Mittwoch)
Pagina: 4
Seite: 3 te von 3
Umfang (Seiten): 3
Zeit: 8:41:11 Uhr